QDR 4500A dual-energy X-ray absorptiometer underestimates fat mass in comparison with criterion methods in adults[1–4]

Dale A Schoeller, Frances A Tylavsky, David J Baer, William C Chumlea, Carrie P Earthman, Thomas Fuerst, Tamara B Harris, Steven B Heymsfield, Mary Horlick, Timothy G Lohman, Henry C Lukaski, John Shepherd, Roger M Siervogel, and Lori G Borrud

ABSTRACT

Background: Dual-energy X-ray absorptiometry (DXA) has become one of the most frequently used methods for estimating human body composition. Although the DXA technique has been validated for the measurement of fat-free mass and fat mass, differences in calibration between instruments produced by different manufacturers, as well as between different models produced by the same manufacturer, have been reported.

Objective: The objective was to compare the calibration of the QDR 4500A against criterion methods in a large heterogeneous population.

Design: DXA-derived body-composition data were obtained from 7 studies: 6 data sets were provided by the investigators, one of which was published. The data included fat mass and fat-free mass measured with a QDR 4500A and criteria measurements of body composition from total body water by dilution at 4 centers, densitometry from 1 center, and four-compartment analysis at 2 centers.

Results: In the cohort of 1195 subjects, 602 men and 593 women aged 19–82 y with a body mass index (in kg/m^2) of 16–44, the fan-beam DXA overestimated fat-free mass ($P < 0.05$). A significant difference was observed in all 7 data sets, and the mean (±SE) was 5 ± 1%.

Conclusions: It is recommended that the lean soft tissue mass estimate with the fan-beam QDR 4500A be reduced by 5% and that for fat mass be increased by that same mass. This finding is particularly important because the National Health and Nutrition Examination Survey is using the QDR 4500A to assess body composition in a nationally representative sample of persons in the United States. *Am J Clin Nutr* 2005;81:1018–25.

KEY WORDS Body composition, hydration, total body water

INTRODUCTION

National data gathered from examination studies over the past 30 y have shown an increase in overweight and obesity across all strata of the US population on the basis of body mass index (BMI; in kg/m^2) (1, 2). Although BMI provides an acceptable approximation of total body fat (3), its use has limits because the relation between BMI and body fat varies with age, sex, physical training, and ethnicity (4–7).

A direct measure of fat-free mass (FFM) and fat mass (FM) is, therefore, often preferred for assessing obesity. Dual-energy X-ray absorptiometry (DXA) estimates of FFM and FM have been validated and generally are reported to correlate highly with values determined with criterion methods (8–14). In addition, DXA can provide estimates of fat distribution by body region, yet is rapid and simple to perform in most subjects. The DXA information on body composition from large multiethnic studies may help to identify factors that might explain differences seen in cardiovascular disease risk factors and other markers of chronic disease, including bone mineral density (BMD) (15–18). For these reasons, the National Center for Health Statistics, Centers for Disease Control and Prevention, included DXA in the current National Health and Nutrition Examination Survey (NHANES) (19).

DXA, however, is not without limitations. Although highly correlated with criterion methods, modest systematic variation in the absolute estimates of body composition by DXA can arise from different hardware and software accommodations to several factors, including interpolations for soft tissues located over bone (8) and treatment of pixels for which a small portion is bone (20). For newer and faster fan-beam instruments, parallax error (beam magnification) due to the variation in heights between the source and detector of tissues is an additional concern (11). Because of these inaccuracies in a given DXA instrument (11), it

[1] From the Nutritional Sciences, University of Wisconsin–Madison (DAS); Preventive Medicine, Health Science Center, University of Tennessee, Memphis, TN (FAT); the Beltsville Human Nutrition Research Center, US Department of Agriculture, Beltsville, MD (DJB); the National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD (LGB); the Lifespan Health Research Center, Department of Community Health, Wright State University School of Medicine, Kettering, OH (WCC and RMS); Virginia Polytechnic Institute and State University, Blacksburg, VA (CPE); the University of California, San Francisco (TF); the National Institute on Aging, National Institutes of Health, Bethesda, MD (TBH); the Obesity Research Center, St Luke's Roosevelt Hospital, New York, NY (SBH and HM); the University of Arizona, Tucson, AZ (TGL); the Grand Forks Human Nutrition Research Center, US Department of Agriculture, Grand Forks, ND (HCL); and Radiology, University of California, San Francisco (JS).

[2] Mention of a trademark or proprietary product does not constitute a guarantee of warranty of the product and does not imply its approval to the exclusion of other products that may also be suitable by any of the following agencies: USDA, CDC/DHHS, or NIH/DHHS.

[3] Supported in part by grants from the National Institute on Aging (N01-AG-2106, N01-AG-6-2102, and N01-AG-6-2103) and by Virginia Polytechnic Biological Initiative Funds. The initial workshop was supported by the National Center for Health Statistics, Centers for Disease Control and Prevention, and the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

[4] Reprints not available. Address correspondence to DA Schoeller, Nutritional Sciences, University of Wisconsin–Madison, 1415 Linden Drive, Madison, WI 53706. E-mail: dschoell@nutrisci.wisc.edu.

Received August 4, 2004.

Accepted for publication December 20, 2004.

TABLE 1
Subject characteristics[1]

Sample	Age	Weight	Height	BMI
	y	*kg*	*cm*	*kg/m²*
1 (n = 0 M, 22 F)	65 ± 4 (59–74)	71 ± 14 (50–104)	165 ± 5 (157–174)	26 ± 5 (19–36)
2 (n = 139 M, 139 F)	75 ± 3 (70–82)	76 ± 16 (41–127)	167 ± 10 (147–192)	27 ± 5 (16–42)
3 (n = 12 M, 12 F)	47 ± 10 (32–66)	78 ± 20 (48–130)	170 ± 11 (147–188)	27 ± 7 (17–42)
4 (n = 68 M, 3 F)	43 ± 11 (19–71)	81 ± 13 (57–117)	164 ± 6 (151–181)	30 ± 3 (22–38)
5 (n = 30 M, 28 F)	74 ± 2 (70–79)	76 ± 15 (38–111)	167 ± 9 (148–187)	27 ± 5 (17–40)
6 (n = 206 M, 245 F)	44 ± 15 (19–79)	78 ± 17 (44–125)	172 ± 10 (148–197)	26 ± 5 (16–44)
7 (n = 147 M, 144 F)	39	63	162	24
Total (n = 602 M, 593 F)	55 ± 15	75 ± 6	167 ± 3	27 ± 1

[1] All values (except totals) are unweighted $\bar{x}$ or $\bar{x}$ ± SD; range in parentheses.

is important to either cross-validate new DXA models or substantive software revisions against an established DXA instrument or an established criterion method.

The accuracy of the QDR 4500A (Hologic, Bedford, MA) is of particular interest because it is being used in the continuous NHANES. The results from NHANES will provide national estimates for body composition, which should be a valuable baseline for future surveys or smaller independent samples. Future studies are likely to use other instruments or techniques; therefore, establishing the accuracy of the fan-beam QDR 4500A is critical. Recent studies, however, have indicated that the Hologic QDR 4500A overestimates FFM and underestimates FM compared with criterion methods (12, 21, 22) and the pencil-beam QDR 2000 (11, 13).

The aim of this study was to compare the QDR 4500A measurement of FM and FFM against that of measurements made with criterion methods at multiple laboratory sites. The use of multiple laboratories ensured that individual bias from any one instrument, criterion method, or protocol would be minimized and, thus, the NHANES results should be the most reliable estimates of body composition in the US population that are possible.

SUBJECTS AND METHODS

Subjects

Seven data sets containing estimates of body composition from the QDR 4500A DXA and a corresponding criterion method were provided from 6 different laboratories. A seventh data set was taken from the literature (22) (**Tables 1** and **2**). The combined cohort ranged in age from 19 to 82 y and in BMI from 16 to 44. During the initial review, 26 data points were dropped from the analysis, including 17 because of a 3-kg weight difference between scale weight and DXA weight, 4 because of a >6-kg difference in FFM between total body water (TBW) and DXA, 3 because they were too large to fit within the DXA detector field, and 2 because TBW accounted for <51% of FFM_{DXA}. The remaining sample included 1004 participants from the 6 data sets and 191 participants from the publication of Deurenberg-Yap et al (22). The human studies were reviewed and approved by the institutional review boards of the respective institutions, and written informed consent was obtained from the participants. All data sets were stripped of participant identifiers before being used in this data analysis.

TABLE 2
Comparison of body-composition data obtained with the Hologic (Bedford, MA) QDR 4500A and with the criterion method[1]

Sample	Criterion method	FFM		FM		Body mass, DXA−Scale
		DXA	DXA − Criterion	DXA	DXA − Criterion	
		kg		*kg*		*kg*
1 (n = 22)	TBW[2]	45.0	2.4 ± 1.4[3]	26.6	−1.7 ± 1.5[3]	0.7[3]
2 (n = 278)	TBW[2]	52.2	4.7 ± 1.9[3]	24.0	−5.1 ± 2.7[3]	−0.2[3]
3 (n = 24)	TBW[2]	55.4	3.0 ± 2.5[3]	23.3	−2.1 ± 2.6[3]	0.9[3]
4 (n = 71)	TBW[2]	50.2	2.0 ± 1.6[3]	30.4	−2.4 ± 1.8[3]	−0.3[3]
5 (n = 58)	4C[4]	53.5	3.3 ± 1.8[3]	23.5	−2.2 ± 1.8[3]	1.1[3]
6 (n = 451)	UWW[5]	54.2	1.8 ± 3.5[3]	23.4	−1.3 ± 3.6[3]	0.5[3]
7 (n = 191)	4C[6]	45.3	1.9[7]	17.5	−1.9[7]	NA

[1] DXA, dual-energy X-ray absorptiometry; TBW, total body water; 4C, 4-compartment model; UWW, underwater weighing; NA, not available.
[2] Assumes hydration of FFM = 0.73.
[3] Significantly different from 0, $P < 0.05$.
[4] Four-compartment model of Withers et al (23).
[5] Densitometry, two-compartment density model of Siri (24).
[6] Four-compartment model of Baumgartner et al (25).
[7] Published analysis showed that DXA underestimated percentage fat.

Protocols

Protocols varied only slightly between laboratories. All participants were measured after an overnight fast and were asked to avoid strenuous exercise on the day before measurement to avoid dehydration. The participants were weighed in a hospital gown or scrubs (data sets 2, 3, 4, and 5) or light clothing, ie, tee shirt and shorts or sweat pants (data sets 1 and 6) and then underwent body-composition analysis with the QDR 4500A and with the criterion method. Further details of the protocols are published elsewhere (12, 21, 26).

Whole-body DXA

A Hologic model QDR 4500A fan-beam X-ray absorptiometer was used to measure body composition at each site. Lean soft tissue mass (LSTM), FM, and bone mineral content (BMC) were assessed by using software versions 8.21, 8.25A, or 8.26A (Whole Body Analysis software), which all use pixel-specific adjustment for beam magnification and the same routine for estimating body composition. Participants were positioned for whole-body scans according to the protocol recommended by the manufacturer. Participants lay supine on the DXA table with limbs close to their bodies. Only those with all body parts in the scan field were included in this study. FFM_{DXA} was the sum of LSTM and BMC.

DXA phantom analysis

To test whether the instruments were comparable, the Hologic whole-body phantom was circulated among laboratories 1 to 5 as well as the 3 NHANES mobile examination centers. At each site, the phantom was scanned 10 times with repositioning of the phantom between each scan; thus, the results included positioning error. The phantom measurements were analyzed and reviewed for accuracy errors by trained staff at the University of California, San Francisco. The mean and SD for the most commonly used variables from each phantom were calculated by using standard methods, including body mass, LSTM, FM, percentage fat, BMD, and BMC. The ratio of LSTM obtained with each instrument to the mean value averaged across all instruments was calculated and compared with the ratio of FFM_{DXA} to $FFM_{criterion}$ for each individual DXA instrument by linear regression.

Criterion methods

Data set 1 was from the Virginia Polytechnic Institute and State University. The criterion method was TBW measured by deuterium dilution. After a baseline venous blood sample was collected, subjects were given a weighed dose of deuterium oxide (99.8 atom%; Isotec Inc, Miamisburg, OH), equivalent to 0.30 g/kg body weight in 100 mL distilled water, followed by a 100-mL distilled water rinse. After a 3-h equilibration period, during which subjects did not ingest anything, a second postdose blood specimen was drawn. Plasma samples were purified by diffusion after the method described by Davis et al (27). Plasma samples were purified by incubating equal volumes of plasma and distilled deionized water at 37 °C for 48 h in incubation dishes (Bel-Air Products, Pequannock, NJ). Purified plasma samples were subsequently analyzed for deuterium enrichment by isotope ratio mass spectrometry (MAT 251; Finnigan, Bremen, Germany). TBW was corrected for urinary deuterium losses and reduced by 4% to correct for exchange with protein and carbohydrate during the 3-h equilibration period (28).

Data set 2 was from the energy expenditure subset of the Health, Aging, and Body Composition (Health ABC) Study conducted at the University of Pittsburgh and the University of Tennessee Health Science Center as described elsewhere (29, 30). The criterion method was TBW measured by isotope dilution; ≈4 g deuterium and 8 g ^{18}O-labeled water were given to fasted subjects after collection of a baseline urine sample. Three additional urine samples were collected in the next 6 h. Enrichments of tracer were measured in the final 2 urine samples relative to baseline by isotope ratio mass spectrometry (Finnigan Delta-S and Delta Plus). Corrections were made for water intake during urine collections, and dilution spaces were reduced by 4% and 0.7% for deuterium and ^{18}O dilution, respectively, and averaged (28).

Data set 3 was from the US Department of Agriculture, Beltsville, MD. The criterion method was TBW measured by isotope dilution. After an overnight fast, subjects provided a urine sample that was used to measure background isotope enrichments and were then dosed orally with 0.1 g/kg body weight each of deuterium and ^{18}O. The bottle containing the dose was rinsed with 100 mL of deionized water and consumed. After that, the subjects were offered a muffin and juice and were allowed coffee, tea, or water for up to 1 h postdose. Food and beverages were not allowed after the first hour. Saliva samples were collected 4 and 4.5 h postdose. Isotopic enrichment was measured by continuous-flow, isotope ratio mass spectroscopy (Europa Scientific Hydra, Cheshire, United Kingdom). TBW was calculated as the average of the deuterium and ^{18}O dilution spaces calculated from the enrichment of the last saliva sample and corrected for 4% and 0.7% in vivo exchange, respectively (28).

Data set 4 was from the University of Tennessee Health Science Center. The criterion method was TBW measured by deuterium dilution as described elsewhere (29, 30). An oral dose of 4 g deuterium oxide was administered to each participant after a 6–12 h fast. Plasma samples were collected into a dry EDTA-coated tube before and 4 h after the isotope administration. Plasma protein was removed by ultrafiltration, and deuterium enrichment above baseline was measured by isotope ratio mass spectrometry (Finnigan Delta Plus). Subjects were allowed fluid at 1 h after the dose, and corrections were made by subtracting this water intake from TBW, but only in those in whom intake exceeded 0.2 kg. A 1% correction was made for isotope exchange during plasma filtration (30) and 4% for in vivo exchange (28).

Data set 5 was from the University of California, San Francisco, as described elsewhere (31). A four-compartment model was used for the criterion method. TBW was measured by using the same methods as described for data set 4. Body density was measured by underwater weighing (UWW) while the subjects wore a bathing suit. Water temperature was set at 32–35 °C. Five of the most consistent trials (underwater weights within a range of 0.02 kg) from 10 replicates were averaged. Before submersion, residual lung volume was measured in triplicate by using a respirometer (model SVR/PLUS; Collins, Braintree, MA). Total-body mineral mass was calculated from measured total-body BMC in the skeleton by the QDR 4500A. To account for the mineral in nonosseous tissue, total mineral from DXA was multiplied by 1.23. The four-compartment model equation of Lohman (32) was used except as noted above, where total-body mineral was used rather than bone mineral (33).

Data set 6 was from the Life Span Health Research Center at the Wright State University School of Medicine. The criterion method was UWW. Underwater weight was measured in a tank of water (4 ft wide, 6 ft long, and 5 ft deep) at 34 °C. The chair was suspended by 4 load cells whose weights were summed and the weight printed. Residual lung volume was measured on land with a SensorMedics model 2450 (Yorba Linda, CA) Pulmonary Function Laboratory. Weight in air was measured on a Seca scale (Hamburg, Germany) to 0.1 kg. The body density data for each adult participant were converted into FFM by using the Siri equation as described in detail by Guo et al (26).

Data set 7 was extracted from published data (22) and included to show external validity. It was the only other study that met our criteria for comparing the QDR 4500A against TBW, UWW, or a four-compartment model as the criterion method. Briefly, these investigators used the four-compartment model of Baumgartner et al (25). Body density was measured by air plethysmography. TBW was measured by deuterium dilution 3 h after a 10-g oral dose of deuterium oxide using plasma. Deuterium was measured by infrared spectroscopy, and a 5% correction for in vivo isotope exchange was applied. Bone mineral was determined by using the QDR 4500A and multiplied by 1.167 to adjust it to the Lunar DPXL equivalent and again by 1.235 to calculate total-body mineral to match the assumptions used in the development of the four-compartment model (25).

Calculation of FFM

When the criterion method was TBW, FFM was calculated as TBW/0.73. FM was calculated as body mass minus FFM. The selection of 0.73 as the hydration of FFM was based on an extensive review of the literature (**Table 3**). When the criterion method was the four-compartment model, percentage fat was calculated by using the equations referenced above. FM was calculated as percentage fat/100 multiplied by body mass, and FFM was calculated as the difference.

Statistical analysis

Means and SDs were calculated by using standard unweighted methods. Within-laboratory means for FFM, FM, and body mass from DXA and the criterion method were compared by using a *t* test. To test for an error in the DXA calibration, regression of FFM_{DXA} on the criterion method results by using the least-squares fit of *y* on *x*. Correlations were identified as significant based on the regression coefficient exceeding the critical value for the given df. To test for a constant offset between methods, the intercept was compared with zero and to test for a proportional error, the slope was compared with unity. To determine the correction factor for $FFM_{DXA,}$ FFM_{DXA} was regressed on $FFM_{criterion}$ while forcing the intercept through zero. In all but 2 of the data sets, the intercept term was not significantly different from zero and thus could be eliminated from the regression. It was eliminated from the other 2 data sets to facilitate comparison between sets. To test for an influence of age, sex, or ethnicity on the DXA_{FFM} correction factor, an analysis of variance for these 3 predictors was performed with control for between-laboratory differences by including laboratory as a predictor. Variances were compared by using the *F* test. Statistical significance required a P value ≤ 0.05. Statistical calculations were performed by using JMP version 4 (SAS Institute Inc, Cary, NC).

RESULTS

The combined data set included 1195 adults who represented a wide range of ages and body sizes (Table 1). The subjects were racially diverse, including 750 whites, 153 African Americans (132 of whom were in sample 2), 1 Asian American, and 291 Asians.

Data were combined for men and women, and the DXA estimate of FFM was compared with that from the criterion method individually for data sets 1–6 (Table 2). FFM was 1.8–4.7 kg larger by DXA than by the criterion method, and FM was 1.3–5.1 kg smaller by DXA in laboratories 1- 6 (Table 3). DXA-determined body mass was significantly different from scale mass in all laboratories, but the differences included both negative and positive offsets, the largest of which was 1.1 kg. Data on FFM and FM were not directly available from sample 7, but it was reported that the DXA-derived percentage fat was 2 and 4 percentage points smaller in women and men, respectively, than that from the four-compartment criterion method. From this we calculated that FFM averaged 1.9 kg greater by DXA and FM 1.9 kg less by DXA.

FFM_{DXA} was found to have a strong linear relation with $FFM_{criterion}$ in each of the data sets. The correlation coefficients ranged from 0.972 in data set 4 to 0.991 in data set 5. When FFM_{DXA} was regressed on $FFM_{criterion}$, the slopes ranged from 0.909 to 0.967 (**Table 4**). The average of the individual slopes was 0.946, which indicates that DXA overestimated FFM by 5.4% with an SE of 0.7% compared with the criterion methods. The SE of the slope for data set indicated that the precision of the linear fit was significantly greater for data sets 1 and 3 than for the other data sets ($P < 0.05$; *F* test). Of note, however, the variance about the slope within a laboratory (V_{WL}) for any given data set is less than the variance in the slopes (V_{tot}) when averaged between data sets ($P < 0.01$; *F* test). Calculation of the between–data set (laboratory) variance ($V_{BL} = V_{total} - \Sigma V_{WL}$) indicates that the between–data set SD for the slope was 0.015. This value is more than twice the largest within–data set SE and, therefore, the between–data set (laboratory) variance has a greater influence on the final average across data sets than does the individual variance within a data set.

As a further test of between-laboratory variance, a phantom was exchanged among laboratories for analysis (**Table 5**). The cross-validation identified small but significant differences between instruments with regard to mass, FM, and BMC. The largest difference between instruments for FFM (LSTM + BMC), however, was only 396 g, or 1.4% of the phantom mass, and thus smaller than the FFM differences between DXA and the criterion methods identified in Table 2. This indicated that the between-laboratory error was probably not due to instrument-to-instrument variation.

We further tested whether the between-laboratory differences observed in the phantom analysis could explain differences in the comparison of FFM_{DXA} and $FFM_{criterion}$. We could not compare the absolute errors from the phantom data with those of the human data because of the differences in total mass between the phantom and the humans. Thus, it was necessary to express the phantom data as the ratio of LSTM for each laboratory to the average LSTM for all laboratories so that it could be compared with the percentage error in

TABLE 3
Review of the literature in which hydration of fat-free mass (FFM) was determined by using the four-compartment model or a model with similar strength[1]

Criterion method	Post hoc correction[2]	Subjects				FFM hydration	Reference
		Race	Sex	Age	Miscellaneous		
				y		%	
IVNA	—	NA	M	22–92	Malnourished	73.2 ± 2.4[3,4]	Beddoe et al (34)
IVNA	—	NA	F	20–88	Malnourished	75.0 ± 3.1[4]	Beddoe et al (34)
IVNA	—	NA	M	20–58		71.1 ± 1.2	Beddoe et al (34)
IVNA	—	NA	F	19–59		72.6 ± 1.5	Beddoe et al (34)
4C	H exchange	W	M	65–94		74.5 ± 4.5	Baumgartner et al (25)
4C	H exchange	W	F	65–94		74.6 ± 3.9	Baumgartner et al (25)
4C	—	NA	M	18–59		73.3 ± 2.2	Fuller et al (35)
4C	—	NA	F	18–59		74.5 ± 1.9	Fuller et al (35)
3C[5]	H exchange	W	M	22–39		70.2 ± 1.0	Hewitt et al (36)
3C[5]	H exchange	W	M	65–85		71.8 ± 1.1	Hewitt et al (36)
3C[5]	H exchange	W	F	22–39		69.9 ± 1.3	Hewitt et al (36)
3C[5]	H exchange	W	F	65–85		71.5 ± 1.7	Hewitt et al (36)
IVNA	—	NA	M	23–72		72.5 ± 0.8	Ryde et al (37)
IVNA	—	NA	F	23–72		72.2 ± 0.8	Ryde et al (37)
4C	H exchange	NA	M and F	30 ± 4		73.3 ± 2.0	Mazariegos et al (38)
4C	H exchange	NA	M and F	74 ± 7		72.7 ± 3.0	Mazariegos et al (38)
4C	H exchange	NA	M	28 ± 4	Runners	73.7 ± 0.8	Penn et al (39)
4C	H exchange	NA	M	28 ± 4		72.9 ± 1.6	Penn et al (39)
3C (Siri)	—	NA	F	—	Gravid (3rd trimester)	76.2	Calalano et al (40)
4C	H exchange	W	M	24 ± 4	Resistance training	74.4 ± 1.2[4]	Modlesky et al (41)
4C	H exchange	W	M	24 ± 4		71.2 ± 2.0	Modlesky et al (41)
4C	H exchange	NA	M and F	19–27		72.5 ± 1.0	Bergsma-Kadijk et al (42)
4C	H exchange	NA	M and F	65–78		73.9 ± 2.5	Bergsma-Kadijk et al (42)
3C (Siri)	—	Asian	M	—		70.4 ± 2.4	Borgounha et al (43)
3C (Siri)	—	Asian	F	—		71.9 ± 2.4	Borgounha et al (43)
4C	—	—	F	30 ± 4	>3 mo postpartum	73.3 ± 2.0	Butte et al (44)
4C	H exchange	W	M	20–94		74.1 ± 3.2	Visser et al (45)
4C	H exchange	AA	M	20–94		74.6 ± 2.8	Visser et al (45)
4C	H exchange	W	M	20–94		73.9 ± 3.4	Visser et al (45)
4C	H exchange	AA	F	20–94		75.3 ± 3.6	Visser et al (45)
4C	—	NA	M	69 ± 7		74.7 ± 3.8	Goran et al (9)
4C	—	NA	F	69 ± 7		72.4 ± 4.6	Goran et al (9)
4C	—	NA	M	26 ± 6	Athletically trained	71.1	Withers et al (23)
4C	—	NA	F	26 ± 6	Athletically trained	70.8	Withers et al (23)
4C	—	NA	M	26 ± 6		70.5	Withers et al (23)
4C	—	NA	F	26 ± 6		71.4	Withers et al (23)
4C	—	NA	M and F	—	Young adult	73.2 ± 2.4	Ritz et al (46)
4C	—	NA	M and F	>60		73.4 ± 2.4	Ritz et al (46)
4C	—	Asian	M	23 ± 4		73.2 ± 1.7	Werkman et al (47)
4C	—	Asian	F	23 ± 4		72.8 ± 1.5	Werkman et al (47)
4C	—	W	M	23 ± 4		72.9 ± 1.9	Werkman et al (47)
4C	—	W	F	23 ± 4		74.2 ± 1.4	Werkman et al (47)

[1] IVNA, in vivo neutron activation; 4C, four-compartment model; 3C, three-compartment model; AA, African American; W, white; NA, not available.
[2] Recalculated from published data to include a correction for hydrogen tracer exchange with nonaqueous material of 4.2% (48).
[3] $\bar{x}$ ± SD (all such values).
[4] Significantly different from the study's identified control group, $P < 0.05$.
[5] Adjusted for between-individual differences in bone mineral density.

the human FFM_{DXA} data for each laboratory. This ratio for the phantom measurement for each laboratory was then regressed onto the ratio of FFM_{DXA} to $FFM_{criterion}$ from those same laboratories. There was no correlation ($r = 0.13$, NS) indicating that instrument calibration as measured by using the phantom did not explain the variation in the ratio of FFM_{DXA} to $FFM_{criterion}$ between laboratories.

Because the criterion method might also influence the comparison of FFM_{DXA} with $FFM_{criterion}$, we compared the slopes of the regression of FFM_{DXA} on $FFM_{criterion}$ using multiple criterion methods for those data sets in which comparisons between several criterion methods could be made (**Table 6**). This was only possible in 2 data sets, however. The slopes and SDs (0.952 ± 0.019) for the various criterion methods from these 2 data sets were not different from those for the between-laboratory variance (0.945 ± 0.016; Table 4).

The subjects were not equally stratified for sex, race, and age across the laboratories. To test whether the slope for the FFM_{DXA} compared with FFM criterion, and hence the correction factor for FFM_{DXA}, differed by sex, age, or ethnicity, we performed an

TABLE 4
Summary of the proportional error in fat-free mass measured by dual-energy X-ray absorptiometry compared with the proportional error in that measured with criterion method based on linear regression analysis[1]

Sample	Criterion	Slope	SE
1 (n = 22)	TBW	0.946[2]	0.006
2 (n = 278)	TBW	0.909[2]	0.002
3 (n = 24)	TBW	0.941[2,3]	0.007
4 (n = 71)	TBW	0.957[2,3]	0.003
5 (n = 58)	4C	0.950[2]	0.003
6 (n = 451)	UWW	0.967[2]	0.003
7 (n = 291)	4C	0.959[2]	0.002
Total	—	0.946[2,4]	0.007[4]

[1] TBW, total body water; 4C, four-compartment model; UWW, underwater weighing.
[2] Significantly greater than 1.0.
[3] Intercept was significantly different from 0 before forcing the regression through 0.
[4] Average and SE of the mean for the slopes for the 7 data sets.

analysis of variance while controlling for laboratory. Sex, age, and ethnicity did not significantly influence the slope.

Given that between-laboratory variance in the relation between FFM from DXA and the criterion methods was larger than the within-laboratory variance and that the variance could not be explained by variance in the analysis of the phantom data, we concluded that the variance resulted from small between-laboratory biases in the criterion methods. Because of this, the average of the mean proportional error in FFM for the 7 data sets was used to estimate the bias in the body-composition estimates of the QDR 4500A. This average was 0.946 and it was significantly different from zero ($P < 0.001$). Thus, the calibration of the QDR 4500A was in error, which resulted in an overestimate of FFM and an underestimate of FM.

A correction factor for QDR 4500A estimates of FFM and FM was determined by using the assumption that BMC obtained from the QDR 4500A is correct and should not be adjusted. Corrected FFM_{DXA} was obtained by multiplying LSTM by the determined correction factor. Corrected FM was then calculated by subtracting the sum of LSTM and BMC from total body weight determined by DXA as follows:

$$LSTM_{corrDXA} = 0.946\ LSTM_{DXA} \quad (1)$$

$$FM_{corrDXA} = DXA\ weight - (0.946\ LSTM_{DXA} + BMC) \quad (2)$$

$$Percentage\ fat = 100(FM_{corrDXA}/weight_{DXA}) \quad (3)$$

BMC was not corrected.

TABLE 6
Comparison of fat-free mass (FFM) measured by dual-energy X-ray absorptiometry (DXA) with FFM measured by different body-composition criterion models[1]

Sample and model	FFM	FM
	kg	*kg*
5 (n = 58)		
DXA[2]	53.5 ± 12.2	23.5 ± 7.4
TBW[3]	50.1 ± 10.7 (0.935)	25.9 ± 8.0
UWW[4]	50.4 ± 12.5 (0.947)	25.4 ± 8.4
4C[5]	50.3 ± 11.5 (0.932)	25.5 ± 8.1
4C[6]	50.9 ± 11.8 (0.953)	25.0 ± 8.0
7[7] (n = 291)		
DXA[2]	45.3	17.5
TBW[3]	43.0 (0.951)	19.8
UWW[4]	44.6 (0.989)	18.1
4C[6]	43.4 (0.959)	19.4

[1] All values are $\bar{x}$ or $\bar{x}$ ± SD; slope of the linear relation ($FFM_{criterion}$ = slope × FFM_{DXA}) in parentheses. TBW, total body water; UWW, underwater weighing; 4C, four-compartment model.
[2] Measured with the Hologic (Bedford, MA) QDR 4500A.
[3] FFM = TBW/0.73.
[4] Siri equation (24).
[5] Model of Withers et al (49).
[6] Model of Baumgartner et al (30).
[7] Calculated from published data of Deurenberg-Yap et al (22).

TABLE 5
Interlaboratory comparison of the Hologic (Bedford, MA) whole-body dual-energy X-ray absorptiometry (DXA) phantom[1]

Data set	Mass[2]	LSTM	Fat mass[2]	Fat	BMD[2]	BMC
	g	*g*	*g*	*%*	*g/cm²*	*g*
1 (n = 10)	29290 ± 50	15510 ± 150	13770 ± 140	47.0 ± 0.4	1.14 ± 0.02	715 ± 5
2a (n = 10)	29500 ± 40	15400 ± 120	14090 ± 130	47.8 ± 0.4	1.17 ± 0.02	730 ± 14
2b, 4 (n = 10)[3]	29300 ± 50	15590 ± 160	13710 ± 130	46.8 ± 0.5	1.17 ± 0.02	734 ± 8
3 (n = 10)	29180 ± 40	15230 ± 130	13960 ± 130	47.8 ± 0.4	1.08 ± 0.02	701 ± 7
5 (n = 10)	29140 ± 40	15460 ± 84	13680 ± 95	46.9 ± 0.3	1.08 ± 0.02	717 ± 10
Average of 1, 2a, 2b, 3, and 5	29280 ± 140	15440 ± 140	13840 ± 180	47.3 ± 0.5	1.13 ± 0.05	719 ± 13
NHANES MEC1[4] (n = 10)	29140 ± 50	15370 ± 90	13770 ± 80	47.3 ± 0.3	1.15 ± 0.02	715 ± 8
NHANES MEC2 (n = 10)	29370 ± 40	15520 ± 90	13850 ± 65	47.2 ± 0.2	1.12 ± 0.02	713 ± 6
NHANES MEC3 (n = 10)	29220 ± 50	15310 ± 180	13910 ± 190	47.6 ± 0.6	1.11 ± 0.01	698 ± 9
Average NHANES	29240 ± 120	15400 ± 110	13850 ± 70	47.3 ± 0.2	1.13 ± 0.02	709 ± 9

[1] All values are $\bar{x}$ ± SD; n = the number of repeat scans. LSTM, lean soft tissue mass; BMD, bone mineral density; BMC, bone mineral content; NHANES, National Health and Nutrition Examination Survey.
[2] Significant difference between instruments, $P < 0.05$ (ANOVA).
[3] Data were generated by using 2 DXA instruments in different cities. Data set 4 used the same DXA unit as was used in data set 2b.
[4] NHANES MEC1–3 are the 3 DXA units that were used for the NHANES data acquisition.

DISCUSSION

We compared body-composition estimates obtained from the 7 Hologic QDR 4500A DXA instruments with those determined with several criterion methods in 7 independent data sets. In each, we found that FFM_{DXA} was larger than that of the criterion method. Because of the consistency of the results across criterion methods, we concluded that FFM, as measured with the QDR 4500A DXA, was significantly overestimated. Regression analysis indicated that the difference from the value determined with the criterion methods was a proportional difference; therefore, a proportional correction factor is recommended.

One potential criticism of our findings was that 4 of the 7 data sets used TBW as the criterion method. The use of TBW requires an assumption of a constant hydration of FFM. Wang et al (49) recently provided a theoretical basis for variation in the hydration of FFM in healthy adults and, in so doing, predicted a maximal range for the individual hydration factors of 0.66 to 0.77. Our review of the literature indicated that group mean values are less variable than are those predicted by Wang et al (49), ranging from 0.70 to 0.75 in healthy adults (Table 1). Our assumed value of 0.73 is at the center of these ranges. Individual studies within this compilation indicate that hydration of FFM is a few percentage points larger during pregnancy, in severe protein malnutrition, and in muscle builders (34, 40, 41); these observations are consistent with the theoretical basis of hydration reviewed by Wang et al (49). After exclusion of these 3 subsets, we performed a meta-analysis to determine whether there were effects of age in adulthood, sex, or ethnicity on hydration of FFM. The average hydration of FFM, excluding the 3 groups listed above with significant increases in hydration, was 72.6 ± 1.4% (between-study SD) with no significant effect of age in adulthood, sex, or ethnicity. These lines of evidence support our use of a hydration factor of 0.73 for our populations.

We also questioned the use of different criterion methods in our analysis. To address this concern, we performed further data analyses in 2 of our data sets that had measurements that permitted the calculation of FFM by several different criterion methods (Table 6). The mean FFM was similar among the various criterion methods, and the slopes between DXA and the criterion method were all <1.0. Moreover, the within-criterion method variances for the slopes were comparable with those for TBW. This provides further support for the use of multiple criterion methods and substantiates the use of TBW as a criterion method for assessing the accuracy of the calibration of the QDR 4500A.

Our use of an unweighted mean to calculate the correction factor for FFM rather than a weighted mean for sample size could also be questioned. Our choice of the unweighted mean was based on the between-laboratory variance being larger than the within-laboratory variance. Both the literature review of the FFM hydration data and the data in Tables 4 and 6 indicate a larger between-laboratory variation than would be predicted from the SE for the within-laboratory mean. This indicates that there was a systematic difference in the criterion methods between laboratories and that this difference became limiting for the interlaboratory comparison. Thus, we used the between-laboratory average for the slope of FFM_{DXA} on $FFM_{criterion}$ without weighting, despite greatly different numbers of participants in the data sets.

It is also possible that the between-laboratory variance may have resulted from between-instrument variance of the DXA instruments. A common whole-body phantom was circulated among the laboratories to determine interlaboratory differences for the assessment of FFM. The results of phantom data analysis showed no difference between the DXA instruments used in this calibration study. Because the circulation of the phantom was done 1–5 y after the human data were collected, it may not represent the actual accuracy of the particular instrument during the collection of the human data. We speculated, however, that the differences that existed between these instruments during the period of data collection were not any greater than those measured during the phantom analysis for the following reasons. Although calibration of a DXA instrument can be altered as a result of a major repair to the DXA hardware or the upgrade of software, none of the laboratories reported any major repair to their DXA instrument nor were there any differences in the software versions used to collect the human or phantom data. It should also be noted that the 3 NHANES DXA instruments being used to collect the national data were found to provide estimates of FFM similar to those obtained by the laboratories used to determine our correction factor (Table 5). This finding supports the use of the above-suggested corrections for the NHANES data.

The error for the FFM_{DXA} value for the QDR 4500A may appear surprising in light of the large number of publications that have found DXA to be an accurate means of assessing FFM and FM in adults. Many of these validations, however, were performed with the use of pencil-beam instruments. The QDR 4500A is a fan-beam instrument and thus involves a correction for beam magnification and a new software routine for converting the X-ray absorption data to body composition. Beam magnification has been shown to influence the measurement of FFM and FM by DXA (13). This and the use of the new software appear to have introduced a modest systematic error in the FFM calibration that requires correction. This error is specific to the QDR 4500A and these software versions. Other fan-bean instruments require independent validation to determine whether they too are subject to systematic bias.

The calibration of the QDR 4500A is of particular importance because it is currently being used to acquire body-composition data for a nationally representative sample of individuals in the United States as part of the NHANES. Because these data may be used to make national policy decisions, comparisons over time within the United States, comparisons between countries, and comparisons of study cohorts with the NHANES sample, an assessment of the accuracy of the QDR 4500A was critical. Based on our findings of a biologically significant bias in the QDR 4500A, this correction was implemented in NHANES 1999–2004 data before its release. We suggest that others who use the QDR 4500A should also use this correction for FFM, FM, and percentage fat.

Technical information regarding the QDR 4500A absorptiometer was provided by Thomas L Kelly of Hologic Inc.

LGB conceived and organized the combined effort of the authors. DAS, FAT, and TGL organized the data and data analysis and prepared the first draft of the manuscript. DJB, WCC, CPE, TF, TBH, SBH, HCL, HM, JS, and RMS provided the data sets and edited the manuscript. None of the authors reported having a conflict of interest.

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999–2000. JAMA 2002;288:1723–7.

2. Ogden CL, Flegal KM, Carroll MD, Johnson CL. Prevalence and trends in overweight among US children and adolescents. JAMA 2002;288: 1728–32.
3. National Institutes of Health, National Heart, Lung, and Blood Institute. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. Washington, DC: HHS, Public Health Service, 1998.
4. Jackson AS, Stanforth PR, Gagnon J, et al. The effect of sex, age and race on estimating percentage body fat from body mass index: The Heritage Family Study. Int J Obes Related Metab Disord 2002;26:789–96.
5. Prentice AM, Jebb SA. Beyond body mass index. Obes Rev 2001;2: 141–7.
6. Deurenberg P, Yap M, van Staveren WA. Body mass index and percent body fat: a meta analysis among different ethnic groups. Int J Obes Relat Metab Disord 1998;22:1164–71.
7. Gallagher D, Visser M, Sepulveda D, Pierson RN, Harris T, Heymsfield SB. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? Am J Epidemiol 1996;143:228–39.
8. Lohman TG. Dual energy X-ray absorptiometry. In: Human body composition. In: Roche AF, Heymsfield SB, Lohman TG, eds. Springfield, IL: Human Kinetics, 1996:63–78.
9. Goran MI, Toth MJ, Poehlman ET. Assessment of research-based body composition techniques in healthy elderly men and women using the 4-compartment model as a criterion method. Int J Obes Relat Metab Disord 1998;22:135–42.
10. Wang J, Thornton JC, Horlick M, Formica C, Wang W, Pierson RN Jr. Dual X-ray absorptiometry in pediatric studies: changing scan modes alters bone and body composition measurements. J Clin Densitometry 1999;2:135–41.
11. Tothill P, Hannan WJ. Wilkinson S. Comparisons between a pencil beam and two fan beam dual energy X-ray absorptiometers used for measuring total body bone and soft tissue. Br J Radiol 2001;74:166–76.
12. Visser MFT, Lang T, Salamone L, Harris TB. Validity of fan-beam dual-energy X-ray absorptiometry for measuring fat-free mass and leg muscle mass. J Appl Physiol 1999;87:1513–20.
13. Ellis K, Shypailo R. Bone mineral and body composition measurements: cross-calibration of pencil-beam and fan-beam dual-energy x-ray absorptiometers. J Bone Miner Res 1998;13:1613–8.
14. Kistorp C, Svendsen O. Body composition analysis by dual energy x-ray absorptiometry in female diabetics differ between manufacturers. Eur J Clin Nutr 1997;51:449–54.
15. Siervogel RM, Wisemantle W, Maynard LM, et al. Serial changes in body composition throughout adulthood and their relationships to changes in lipid and lipoprotein levels. The Fels Longitudinal Study. Arterioscler Thromb Vasc Biol 1998;18:1759–64.
16. Daniels SR, Morrison JA, Sprecher DL, Khoury P, Kimball TR. Association of body fat distribution and cardiovascular risk factors in children and adolescents. Circulation 1999;99:541–5.
17. Ryan AS, Berman DM, Nicklas BJ, et al. Plasma adiponectin and leptin levels, body composition, and glucose utilization in adult women with wide ranges of age and obesity. Diabetes Care 2003;26:2383–8.
18. Farmer ME, Harris T, Madans JH, Wallace RB, Cornon1-Huntley J, White LR. Anthropometric indicators and hip fracture. The NHANES I epidemiologic follow-up study. J Am Geriatr Soc 1989;37:9–16.
19. Centers for Disease Control and Prevention. NHANES 1999-2000 public data release file documentation. Internet: http://www.cdc.gov/nchs/data/nhanes/gendoc.pdf (accessed 3 March 2005).
20. Kelly TL, Berger N, Richardson TL. DXA body composition: theory and practice. Appl Radiat Isot 1998;49:511–3.
21. Tylavsky F, Lohman T, Blunt BA, et al. The QDR 4500A dual energy x-ray absorptiometer overestimates fat-free mass compared to criterion methods. J Appl Physiol 2003;94:959–65.
22. Deurenberg-Yap M, Schmidt G, van Staveren WA, Hautvast JG, Deurenberg P. Body fat measurement among Singaporean Chinese, Malays and Indians: a comparative study using a four-compartment model and different two-compartment models. Br J Nutr 2001;85:491–8.
23. Withers RT, LaForge J, Pillans RK, et al. Comparisons of two-, three- and four-compartment models of body composition analysis in men and women. J Appl Physiol 1998;85:238–45.
24. Siri WE. The gross composition of the body. In: Advances in biological and medical physics. In: Tobias CA, Lawrence JH, eds. New York, NY: Academic Press, 1956:239–80.
25. Baumgartner RN, Heymsfield SB, Lichtman S, Wang J, Pierson RN. Body composition in elderly people: effect of criterion estimates on predictive equations. Am J Clin Nutr 1991;53:1345–53.
26. Guo S, Chumlea WC, Roche AF, Siervogel RM. Age- and maturity-related changes in body composition during adolescence into adulthood: The Fels Longitudinal Study. Int J Obes Relat Metab Disord 1997;21:1167–75.
27. Davis JM, Lamb DR, Burgess WA, Bartoli WP. Accumulation of deuterium oxide in body fluids after ingestion of D_2O-labeled beverages. J Appl Physiol 1987;63:2060–6.
28. Racette SB, Schoeller DA, Luke AH, Shay K, Hnilicka JH, Kushner RF. Relative dilution spaces of ^2H- and ^{18}O-labeled water in humans. Am J Physiol 1994;267:E585–90.
29. Tylavsky FA, Lohman TG, Dockrell M, et al. Comparison of the effectiveness of 2 dual-energy X-ray absorptiometers with that of total body water and computed tomography in assessing changes in body composition during weight change. Am J Clin Nutr 2003;77:356–63.
30. Blanc S, Colligan AS, Trabulsi J, et al. Influence of delayed isotopic equilibration on the accuracy of doubly labeled water in the elderly. J Appl Physiol 2002;92:1036–44.
31. Salamone LM, Fuerst T, Visser M, et al. Measurement of fat mass using fan beam dual-energy x-ray absorptiometry: a validation study in elderly adults. J Appl Physiol 2000;89:345–52.
32. Lohman TG. Advances in body composition assessment. Springfield, IL: Human Kinetics, 1992:22.
33. Bartok-Olson CJ, Schoeller DA, Sullivan JC, Clark RR. The "B" in the Selinger Four-Compartment body composition formula should be body mineral instead of bone mineral. Fifth international symposium on in vivo body composition studies. Ann N Y Acad Sci 2000;904:342–4.
34. Beddoe AH, Streat SJ, Hill GL. Hydration of fat-free body in protein depleted patients. Am J Physiol 1985;249:E227–33.
35. Fuller NJ, Jebb SA, Laskey MA, Coward WA, Elia M. Four-compartment model for the assessment of body composition in humans. Clin Sci 1992;82:687–93.
36. Hewitt MJ, Going SB, Williams DP, Lohman TG. Hydration of fat-free mass in children and adults: implications for body composition analysis. Am J Physiol 1993;265:E88–95.
37. Ryde SJ, Birks JL, Morgan WD, Evans CJ, Dutton J. A five-compartment model of body composition of healthy subjects assessed using in vivo neutron activation analysis. Eur J Clin Nutr 1993;47:863–74.
38. Mazariegos M, Wang Z-M, Gallagher D, et al. Differences between young and old females in the five levels of body composition and their relevance to the two-compartment chemical model. J Gerontol 1994;49:M201–8.
39. Penn I-W, Wang Z-M, Buhl KM, Allison DB, Burastero SE, Heymsfield SB. Body composition and two-compartment model assumptions in male long distance runners. Med Sci Sports Exerc 1994;26:392–7.
40. Catalano PM, Wong WW, Drago NM, Amini SB. Estimating body composition in late gestation: a new hydration constant for body density and total body water. Am J Obstet Gynecol 1995;172:1464–70.
41. Modlesky KJ, Lewis RD, Prior BM, Sloniger MA, Rowe DA. Density of fat-free mass and estimates of body composition in male weight trainers. J Appl Physiol 1996;80:2085–96.
42. Bergsma-Kadijk JA, Baumeister B, Deurenberg P. Measurement of body fat in young and elderly women: comparison between a four-compartment model and widely used reference methods. Br J Nutr 1996;75:649–57.
43. Borgonha S, Kuriyan R, Shetty P, Ferro-Luzzi A, Kurpad AV. Body composition by a three compartment model in adult Indian male and female subjects. Indian J Physiol Pharmacol 1997;41:227–33.
44. Butte NF, Hopkinson JM, Ellis KJ, Wong WW, Smith EO. Changes in fat-free mass and fat mass in postpartum women: a comparison of body composition models. Int J Obes Relat Metab Disord 1997;21:874–80.
45. Visser M, Gallagher D, Deurenberg P, Wang J, Pierson RN, Heymsfield SB. Density of fat-free mass: relationship with race, age and level of body fatness. Am J Physiol 1997;272:E781–7.
46. Ritz P. Body water spaces and cellular hydration during healthy aging. Ann N Y Acad Sci 2000;904:474–83.
47. Werkman A, Deurenberg-Yap M, Schmidt G, Deurenberg P. A comparison between the composition and density of the fat-free mass of young adult Singaporean Chinese and Dutch Caucasians. Ann Nutr Metab 2000;44:235–42.
48. Wang Z, Deurenberg P, Wang W, Pietrobelli A, Baumgartner RN, Heymsfield SB. Hydration of fat-free body mass: new physiological modeling approach. Am J Physiol 1999;276:E995–1003.
50. Schoeller DA, Hnilicka JM. Reliability of the doubly-labeled water method for the measurement of total daily energy expenditure in free-living subjects. J Nutr 1996;126:348S–54S.

The American Journal of Clinical Nutrition